

# Subsídios para o estudo da percentagem de carrasca no pinheiro bravo

POR

**ANTÓNIO DE CAMPOS ANDRADA**

Engenheiro silvicultor

## INTRODUÇÃO

Se o problema da casca, para alguns pinheiros, como o *Pinus silvestris*, L., tem uma importância reduzida, o mesmo não acontece com o *Pinus pinaster*, Sol., pinheiro essencialmente carrascudo, que por vezes chega a apresentar 50% de casca, mesmo em árvores com mais de 50 anos.

O facto de ser muito variável, de local para local, o valor da percentagem de casca e também muito grande a amplitude da variação ocasiona cometerem-se, na prática, erros grosseiros, quando se atribue um valor médio à percentagem de casca na avaliação do volume de madeira.

Este assunto carece de ser estudado, pelo pouco que tem sido abordado e pela importância que tem na prática.

São muito numerosos os factores que interveem na variação da percentagem de casca. Podemos agrupá-los em duas categorias: os de natureza inerente à árvore e os de natureza ecológica.

Entre os primeiros há a considerar os caracteres morfológicos da árvore e a idade, relativamente fáceis de estudar, e, possivelmente, outros ainda, de natureza genética, duma investigação menos praticável nas essências florestais, em razão do elevado número de anos necessário para conseguir um resultado.

Entre os segundos, cujos efeitos só indirectamente podem ser apreciados, devem ter uma acção importante o clima e a profundidade

do solo. Com efeito, é um facto observado que, não intervindo outros factores de maior influência, os pinheiros teem maior percentagem de carrasca nos terrenos secos do que nos terrenos húmidos e que os pinheiros mais próximos do mar, desabrigados dos ventos, apresentam crescimentos em altura e em diâmetro mais fracos, sendo, pelo contrário, grande a espessura de carrasca.

Ao atravessarmos uma mancha de pinhal em solo de surraipa, são bem evidentes os pequenos crescimentos anuais, quer em altura, quer em diâmetro, que contrastam com a grande espessura de carrasca.

As características morfológicas da árvore estão de tal modo relacionadas com os factores ecológicos que, ao estudarmos a sua influência na variação da percentagem de casca, apreciamos indirectamente, em parte, a influência destes factores.

O mesmo haveria a dizer, e com mais forte razão, quanto às influências de natureza cultural.

É a análise das variações da percentagem de carrasca, com a idade e com alguns factores morfológicos da árvore, tratada pelos métodos estatísticos, o assunto do presente trabalho.

Os dados para este estudo foram colhidos em 198 árvores abatidas na Mata Nacional de Foja no ano de 1935, quando da abertura da nova rede divisional.

De tal facto resulta ter o exame incidido sobre árvores de todas as idades e sujeitas a condições locais muito diferentes. Esta circunstância, a-pesar-de vir dificultar a análise da correlação da percentagem de casca com os caracteres morfológicos da árvore, não deixa de ser vantajosa, pois dela resulta ficar a população a estudar devidamente representada pelos individuos observados e, portanto, poder-se generalizar a toda a mata a aplicação da tabela de percentagem de casca que pretendemos elaborar.

Nas 198 árvores verificámos uma média de 39,2 para a percentagem de casca, valor elevado que se justifica pelo facto de predominarem as árvores novas.

Na Mata Nacional de Leiria o engenheiro silvicultor J. C. Freire Temudo encontrou uma média de 31% em 1455 árvores, onde predominavam também os pinheiros novos.

Dum estudo comparativo relativo a estas duas matas verificámos igualmente, entre grupos de árvores da mesma idade, serem menos carrascudos os pinheiros da Mata Nacional de Leiria.

Na Mata do Valado as percentagens de carrasca são também mais elevadas do que nesta Mata.

Esta disparidade de resultados mostra-nos como é grande a influência das condições locais (especialmente a profundidade do solo) e, possivelmente, o modo de tratamento, na percentagem de carrasca, o que vem dificultar consideravelmente o problema.

Foram os seguintes os caracteres que serviram de base ao nosso estudo (1):

$X_1$  — Percentagem de casca.

$X_2$  — D. A. P.

$X_3$  — Altura total da árvore.

$X_4$  — Idade.

$X_5$  — Largura da copa.

$X_6$  — Percentagem da espessura de casca, em relação ao D. A. P. sobre casca.

Na análise que fizemos, pelos métodos estatísticos, das correlações existentes entre estes factores, não pretendemos atingir conclusões de carácter definitivo.

Seria interessante determinar os limites mínimos de percentagem de carrasca, relacionados com as condições individuais e do meio, e procurar o mais vantajoso processo de os atingirmos, não deixando de ter em vista o tipo de explorabilidade adoptado. Mas o pouco tempo disponível não nos permitiu encarar este aspecto do problema; é extremamente ingrato trabalhar em estatística, pois só à custa de muitas tentativas e de muito trabalho perdido, que não transparece no estudo apresentado, se consegue alguma coisa de útil.

Dividimos a matéria em duas partes: na primeira estudámos a variação da percentagem de carrasca com diversos caracteres inerentes à árvore, que procurámos fôsem no maior número possível, com o fim de conseguir um valor elevado para o coeficiente de determinação; na segunda, com o objectivo de dar uma finalidade prática ao nosso trabalho, organizámos uma tabela de percentagem de carrasca, baseando-nos na análise de variância para a escolha dos factores a considerar. Foi nesta parte que mais dificuldades encontrámos. A solução que adoptámos, ainda que tenha certos inconvenientes de ordem prática, pareceu-nos a mais aconselhável; na altura devida faremos as necessárias considerações justificativas.

---

(1) Todos os elementos numéricos que serviram de base ao presente trabalho foram reunidos no quadro I.

## I PARTE

### Aplicação dos métodos de correlação ao estudo da variação da percentagem de casca

---

#### Considerações gerais

É um facto conhecido que em pinheiros velhos a percentagem de carrasca é menor do que nos pinheiros novos, que em árvores da mesma idade a percentagem de carrasca diminue com a altura e o diâmetro, e, ainda, que ela é maior nas árvores em maciço do que nas árvores vivendo isoladamente.

Havendo uma forte correlação entre estes caracteres, é de grande interesse avaliar não só o modo como cada um actua, mas também a parte da variação da percentagem de carrasca, que cada factor pode explicar, sem que nenhum dos outros factores nela intervenham, e ainda a influência global de todos eles, considerados conjuntamente.

O problema é duma grande complexidade, pois influem sempre numerosos factores, alguns impossíveis de determinar directamente, e outros que, pela sua intervenção conjunta, encobrem a verdadeira relação entre qualquer deles e a variável dependente.

Desconhecendo a importância dos diversos factores, começámos por ensaiar os mais facilmente determináveis, diâmetro e altura total da árvore, e verificámos que a percentagem das causas de variação da percentagem de casca, relativas a esses factores, era apenas de 48,4 0/0. Introduzimos, por isso, outros factores, que fomos ensaiando na sequência da investigação, de modo a obter um valor mais elevado para o coeficiente de determinação.

Para facilitar a exposição da matéria e a elaboração dos quadros,

resolvemos analisar simultaneamente os diversos caracteres em que baseámos o nosso trabalho, como se eles tivessem sido previamente conhecidos.

### Coefficientes de correlação simples

Como ponto de partida para estudos subseqüentes, calculámos os coeficientes de correlação simples pelo emprêgo da fórmula de Pearson (1):

$$r = \frac{\sum (xy)}{\sqrt{(\sum x^2)(\sum y^2)}} \quad (1)$$

No quadro A inscreveram-se os dez valores dos coeficientes de correlação simples.

### QUADRO A

#### Coefficientes de correlação simples

	1	2	3	4
5	— 0.670	0.907	0.787	0.774
4	— 0.525	0.850	0.855	
3	— 0.695	0.894		
2	— 0.628			

1 = casca (%).

2 = D. A. P.

3 = altura total.

4 = idade.

5 = largura da copa.

### Equações de regressão simples e sua representação gráfica

Para avaliar o modo como a percentagem de carrasca varia com cada uma das variáveis independentes, determinámos as linhas de regressão simples. Admitimos, em primeira análise, que as equações



de regressão entre a variável dependente e cada uma das variáveis independentes são do tipo:  $Y = a + b X$ , isto é, que são rectilíneas as linhas de regressão.

Os coeficientes de regressão foram calculados pela fórmula:

$$b = r \frac{\sigma_y}{\sigma_x} \text{ e a ordenada na origem pela fórmula: } a = M_y - b \cdot M_x.$$

As equações de regressão obtidas foram as seguintes:

$$X_1 = 48.828 - 0.476 X_3$$

$$X_1 = 55.616 - 1.180 X_3$$

$$X_1 = 47.502 - 0.179 X_4$$

$$X_1 = 47.605 - 2.988 X_5$$

Nos gráficos 1, 2, 3 e 4 foi feita a representação gráfica destas equações. Para se poder fazer uma ideia da natureza linear ou curvilínea da relação entre a variável dependente e cada uma das variáveis independentes, traçaram-se também nesses gráficos as linhas das médias das classes.

O exame dos gráficos leva-nos a tirar as seguintes conclusões:

1.<sup>a</sup> — A relação entre o D. A. P. e a percentagem de casca é muito ligeiramente curvilínea.

2.<sup>a</sup> — A relação entre a altura e a percentagem de casca é sensivelmente rectilínea.

3.<sup>a</sup> — Para as idades, a linha das médias é bastante confusa, e apenas nos indica que a percentagem de casca varia muito lentamente com a idade dos 15 aos 45 anos, decrescendo depois bruscamente dos 45 aos 75, para voltar a manter-se quasi estacionária dos 75 anos em diante.

4.<sup>a</sup> — A relação entre a largura da copa e a percentagem de casca parece ser curvilínea, devendo ser parabólica a curva de regressão.

5.<sup>a</sup> — Aparentemente a percentagem de casca varia na razão inversa do D. A. P. ( $X_2$ ), da altura ( $X_3$ ), da idade ( $X_4$ ) e da largura da copa ( $X_5$ ).

### Equações de regressão múltipla (três variáveis)

Os coeficientes de correlação simples (quadro A, coluna 1) mostram que nenhuma das variáveis independentes consideradas pode, só por si, ser utilizada para a avaliação da percentagem de carrasca. Tor-

GRÁFICO 1

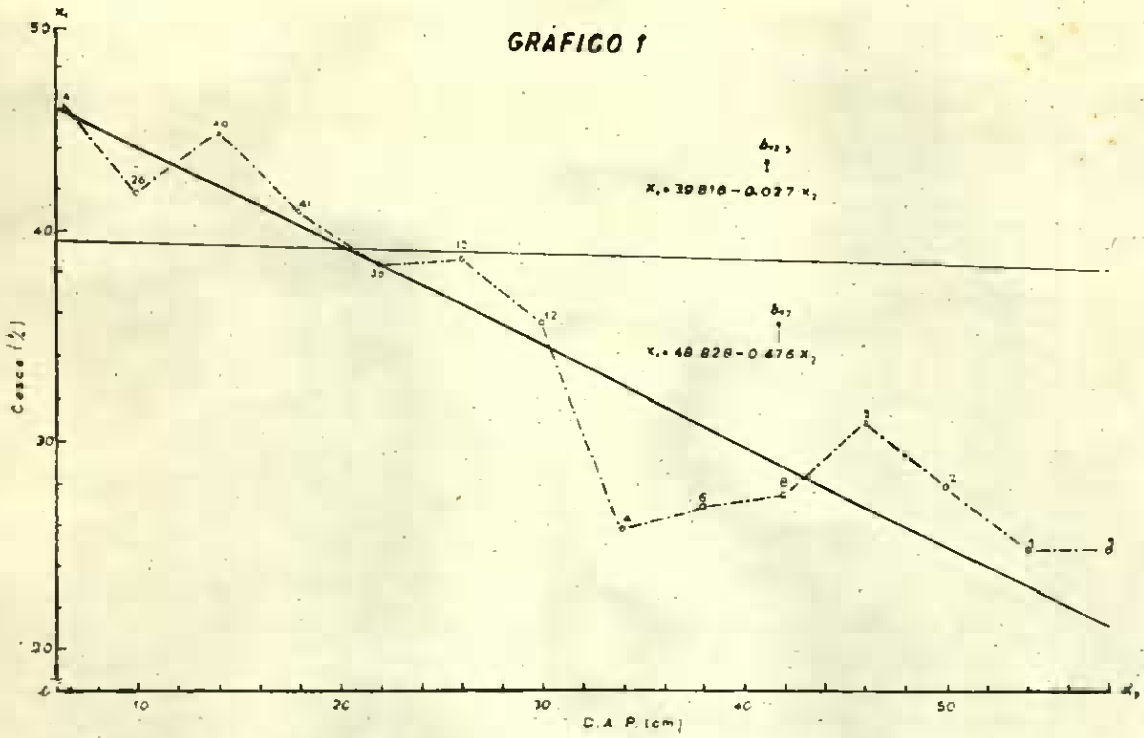
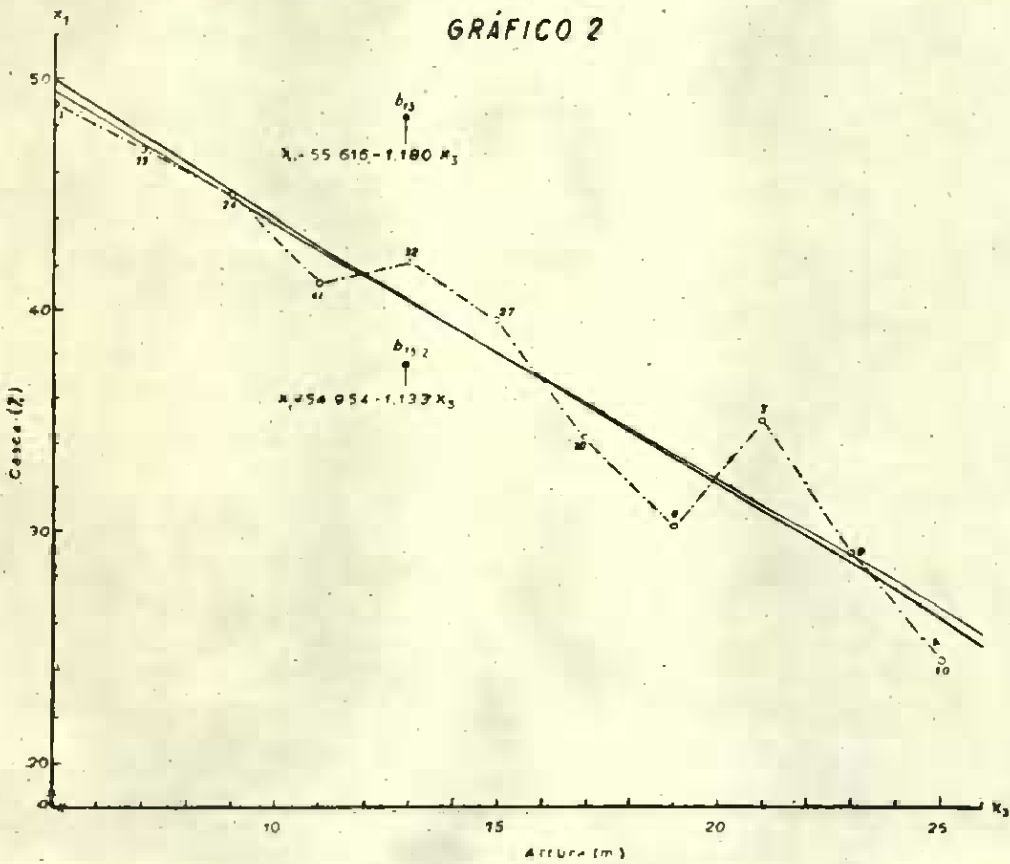
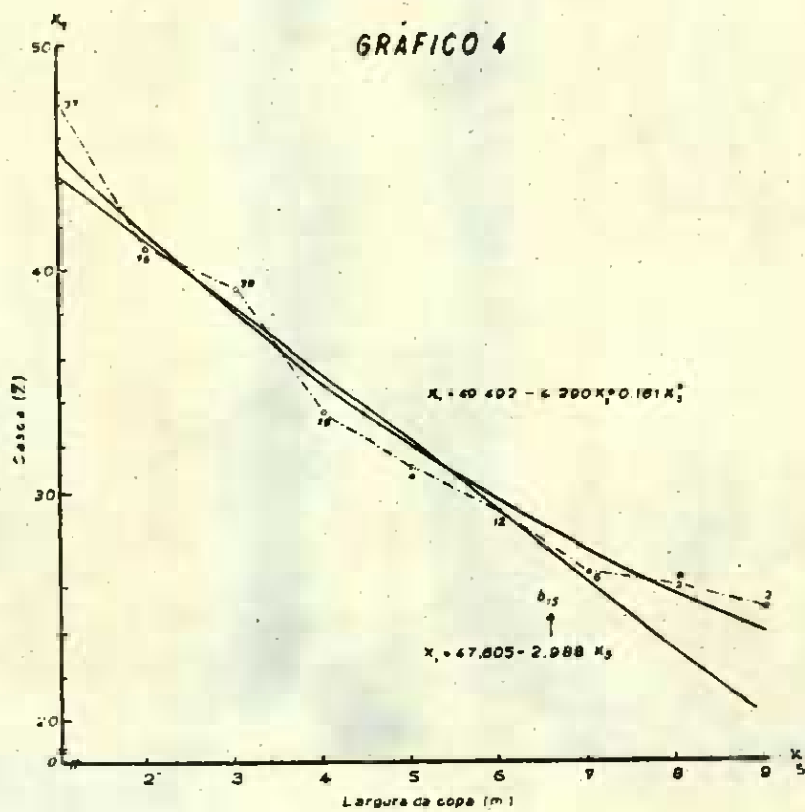
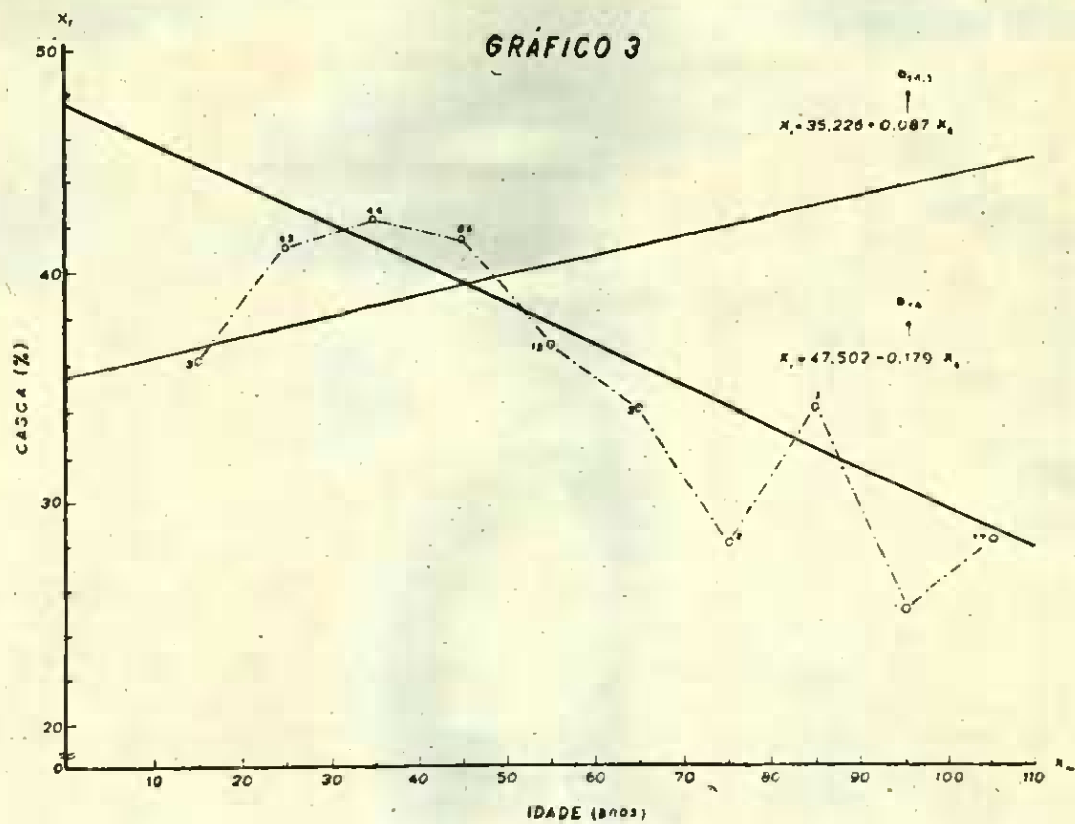


GRÁFICO 2







na-se, por isso, necessário agrupá-las duas a duas, para examinar as relações entre a percentagem de casca e os factores de cada grupo e investigar o que fôr possível acêrca da sua interdependência, continuando para isso a recorrer aos métodos baseados no princípio da correlação rectilínea.

Com este objectivo, ajustámos um plano aos valores observados de cada grupo de três variáveis, mediante as seis equações de regressão:

$$X_1 = a + b_{12.3} X_2 + b_{13.2} X_3$$

$$X_1 = a + b_{12.4} X_2 + b_{14.2} X_4$$

$$X_1 = a + b_{12.5} X_2 + b_{15.2} X_5$$

$$X_1 = a + b_{13.4} X_3 + b_{14.3} X_4$$

$$X_1 = a + b_{13.5} X_3 + b_{15.3} X_5$$

$$X_1 = a + b_{14.5} X_4 + b_{15.4} X_5$$

Para o ajustamento seguiu-se o método dos mínimos quadrados. As equações normais, expressas em unidades padrões:

$$\begin{cases} r_{12} = \beta_{12.3} + \beta_{13.2} r_{23} \\ r_{13} = \beta_{13.2} + \beta_{12.3} r_{12} \end{cases}$$

permitem-nos obter os valores dos coeficientes de regressão ( $\beta$ ):

$$\beta_{12.3} = \frac{r_{12} - r_{23} \cdot r_{13}}{1 - r_{23}^2} \quad \text{e} \quad \beta_{13.2} = \frac{r_{13} - r_{23} \cdot r_{12}}{1 - r_{23}^2} \quad (2)$$

Passa-se dos coeficientes  $\beta$  para os coeficientes  $b$  pelas fórmulas:

$$b_{12.3} = \frac{\sigma_1}{\sigma_2} \beta_{12.3} \quad \text{e} \quad b_{13.2} = \frac{\sigma_1}{\sigma_3} \beta_{13.2} \quad (3)$$

Os coeficientes  $a$  (ordenada na origem) são dados por:

$$a = M_{X_1} - M_{X_2} b_{12.3} - M_{X_3} b_{13.2}$$

Aplicando estas fórmulas, calcularam-se os coeficientes de regressão, que se inscreveram no quadro B.

Depois de decodificar os coeficientes, obtivemos as seguintes equações de regressão:

$$\begin{aligned}
 X_1 &= 55.528 - 0.027 X_2 - 1.133 X_3 & (a) \\
 X_1 &= 49.258 - 0.496 X_2 + 0.011 X_3 & (b) \\
 X_1 &= 48.100 - 0.085 X_2 - 2.520 X_3 & (c) \\
 X_1 &= 56.777 - 1.554 X_2 + 0.087 X_3 & (d) \\
 X_1 &= 53.678 - 0.750 X_2 - 1.440 X_3 & (e) \\
 X_1 &= 47.685 - 0.005 X_2 - 2.934 X_3 & (f)
 \end{aligned}
 \tag{4}$$

### Coefficientes de correlação múltipla e de correlação parcial

Como temos em vista a escolha dos factores de maior influência na percentagem de casca, torna-se necessário conhecer não só a percentagem total das causas de variação adstritas às variáveis independentes, mas também ajuizar o valor relativo de cada um dos caracteres ensaiados, não como caracteres isolados, mas, pelo contrário, tomando-se em consideração a acção recíproca que eles possam exercer entre si.

Considerando primitivamente as variáveis independentes agrupadas duas a duas, foram calculados, com este objectivo, os coeficientes de correlação total e parcial.

Para os primeiros empregou-se a fórmula geral:

$$R_{1,2,3 \dots n}^2 = \frac{\left\{ b_{12,3 \dots n} (\sum X_1 X_2) + b_{13,2 \dots n} (\sum X_1 X_3) + \dots + b_{1n,2,3 \dots (n-1)} (\sum X_1 X_n) \right\}}{\sum (x_1^2)} \tag{5}$$

e para os segundos a fórmula:

$$r_{12,n} = \frac{r_{12} - r_{1n} \cdot r_{2n}}{\sqrt{(1 - r_{1n}^2)} \sqrt{(1 - r_{2n}^2)}} \tag{6}$$

Quando é pequeno o número de observações e grande o número de variáveis, os valores dados pela fórmula (5) devem ser corrigidos, pois tendem a aumentar. O valor mais provável,  $\bar{R}_{1,2,3 \dots n}$  é dado pela fórmula:

$$\bar{R}_{1,2,3 \dots n}^2 = 1 - (1 - R_{1,2,3 \dots n}^2) \left( \frac{n-1}{n-m} \right) \tag{5, a}$$

sendo  $n$  o número de observações e  $m$  o número de constantes da equação de regressão.

No nosso trabalho não tivemos necessidade de aplicar a fórmula (5, a), pois verificámos serem insignificantes as diferenças dos valores dados pelas duas formulas.

O coeficiente de determinação total (quadrado do coeficiente de correlação múltipla) indica-nos a fracção da variação linear da percentagem de casca que se pode atribuir às variáveis independentes consideradas, mas não nos explica o modo como cada uma intervém.

Recorremos, para isso, aos coeficientes de determinação parcial (valores dados pela fórmula (6) elevados ao quadrado); estes valores permitem-nos avaliar a parte da variação da percentagem de casca, que é devida exclusivamente a um dado factor, tomada relativamente à variação que ficou por explicar pelos outros factores considerados simultaneamente.

O sinal positivo ou negativo dos coeficientes de correlação parcial indica-nos ainda o sentido da variação da percentagem de cascas com cada um dos factores ensaiados, tomando em consideração a acção recíproca existente entre eles.

No quadro B foram inscritos os valores dos coeficientes de correlação múltipla, determinação total, correlação parcial, determinação parcial e de regressão.

### **Apreciação dos resultados obtidos**

Se a análise numérica permite pormenorizar mais as conclusões a tirar, o exame dos gráficos tem a vantagem de dar uma visão nítida dessas conclusões. Com este objectivo traçámos as linhas de regressão parciais para cada uma das equações (4).

Substituindo nas equações de regressão, alternadamente, uma das variáveis independentes pelos valores das suas médias aritméticas, obtivemos as seguintes equações de regressão parciais:

Da equação (a),

$$X_1 = 54.954 - 1.133 X_2$$

$$X_1 = 39.816 - 0.027 X_2$$

Da equação (b),

$$X_1 = 38.757 + 0.011 X_2$$

$$X_1 = 49.765 - 0.496 X_2$$

## QUADRO B

## COEFICIENTE

correlação múltipla	determinação total	correlação parcial	determinação parcial	regressão (b)	regressão (b)
$R_{1-23} = 0.697$	$R_{1-23}^2 = 0.486$	$r_{12-3} = -0.022$ $r_{13-2} = -0.384$	$r_{12-3}^2 = 0.0005$ $r_{13-2}^2 = 0.1475$	$\beta_{12-3} = -0.035$ $\beta_{13-2} = -0.667$	$b_{12-3} = -0.035$ $b_{13-2} = -0.756$
$R_{1-24} = 0.629$	$R_{1-24}^2 = 0.395$	$r_{12-4} = -0.406$ $r_{14-2} = 0.022$	$r_{12-4}^2 = 0.1648$ $r_{14-2}^2 = 0.0005$	$\beta_{12-4} = -0.655$ $\beta_{14-2} = 0.032$	$b_{12-4} = -0.662$ $b_{14-2} = 0.036$
$R_{1-25} = 0.670$	$R_{1-25}^2 = 0.449$	$r_{12-5} = -0.064$ $r_{15-2} = -0.306$	$r_{12-5}^2 = 0.0041$ $r_{15-2}^2 = 0.0936$	$\beta_{12-5} = -0.113$ $\beta_{15-2} = -0.565$	$b_{12-5} = -0.114$ $b_{15-2} = -0.840$
$R_{1-31} = 0.709$	$R_{1-31}^2 = 0.502$	$r_{12-1} = -0.558$ $r_{11-2} = 0.185$	$r_{12-1}^2 = 0.3114$ $r_{11-2}^2 = 0.0342$	$\beta_{12-1} = -0.914$ $\beta_{11-2} = 0.256$	$b_{12-1} = -1.036$ $b_{11-2} = 0.291$
$R_{1-35} = 0.723$	$R_{1-35}^2 = 0.523$	$r_{13-5} = -0.367$ $r_{15-3} = -0.277$	$r_{13-5}^2 = 0.1347$ $r_{15-3}^2 = 0.0770$	$\beta_{13-5} = -0.441$ $\beta_{15-3} = -0.323$	$b_{13-5} = -0.500$ $b_{15-3} = -0.480$
$R_{1-45} = 0.669$	$R_{1-45}^2 = 0.448$	$r_{11-5} = -0.013$ $r_{15-1} = -0.491$	$r_{11-5}^2 = 0.0002$ $r_{15-1}^2 = 0.2411$	$\beta_{11-5} = -0.015$ $\beta_{15-1} = -0.658$	$b_{11-5} = -0.017$ $b_{15-1} = -0.978$

Da equação (c),

$$X_1 = 46.300 - 2.520 X_2$$

$$X_1 = 41.049 - 0.085 X_2$$

Da equação (d),

$$X_1 = 35.226 + 0.087 X_2$$

$$X_1 = 60.789 - 1.554 X_2$$

Da equação (e),

$$X_1 = 43.277 - 1.440 X_2$$

$$X_1 = 49.649 - 0.750 X_2$$

Da equação (f),

$$X_1 = 47.454 - 2.934 X_2$$

$$X_1 = 39.476 - 0.005 X_2$$

Fizemos a representação gráfica de todas estas equações, mas só reproduzimos nos gráficos 1, 2, 3 e 4 as que mais interessavam.

O exame destes gráficos e a discussão dos valores inscritos no quadro B permitem-nos tirar as seguintes conclusões:

1.<sup>a</sup> — Do valor de  $R^2_{1.23}$  concluímos que 48,6% das variações da percentagem de casca são devidas ao diâmetro e à altura; por outro lado  $r^2_{12.3}$  e  $r^2_{13.2}$  dizem-nos que o diâmetro explica apenas 0,05% da variação, deixada de explicar pela altura, ao passo que esta permite avaliar 14,75% das variações de percentagem de casca, em que o diâmetro não intervém. O exame dos gráficos 1 e 2 mostra-nos também como é pequena a influência do diâmetro. A linha de regressão parcial  $X_1 = 39.816 - 0.027 X_2$  apresenta-se quase horizontal, com um coeficiente angular muito próximo de zero ( $b_{12.3} = -0.027$ ); isto significa que, tomando em consideração a influência que é devida à altura, fixada esta na sua média, a percentagem de casca quase não varia com o diâmetro. Pelo contrário, o grande ângulo que a linha de regressão simples forma com a linha de regressão parcial patenteia a importância da altura na variação da percentagem de casca.

É o que se conclue também do exame do gráfico 2: o pequeno ângulo, que as duas linhas de regressão formam entre si, resultante da pequena diferença existente entre o coeficiente de regressão simples

$b_{12} = -1.180$  e o coeficiente de regressão parcial  $b_{12.3} = -1.133$ , permite avaliar a pequena influência do diâmetro.

Podemos, portanto, em última análise afirmar que:

a) a elevada correlação entre o diâmetro e a altura ( $r_{23} = 0.894$ ) encobre a verdadeira relação existente entre a percentagem de carrasca e o diâmetro;

b) desde que se tome em consideração a altura, o diâmetro influencia tão pouco na variação da percentagem de casca, que não vale a pena considerá-lo para a organização duma tabela de percentagem de carrasca, apesar de ser um elemento fácil de medir;

c) pelo contrário, a altura apresenta-se como um elemento de avaliação do maior valor, por representar 48,30 % ( $r^2_{13} = 0.483$ ) das causas de variação da percentagem de carrasca e ser fácil a sua determinação.

2.<sup>a</sup> — Como o coeficiente de determinação para a altura e idade é  $R^2_{13.4} = 0.502$ , podemos concluir desde já que, não sendo muito apreciável a influência da idade, ela é mais importante do que a do diâmetro;  $r^2_{14.3} = 0.034$  dá-nos a percentagem da variação da percentagem de casca, devida à idade e que a altura não pode explicar.

O valor positivo de  $r_{14.3}$  mostra-nos que a percentagem de casca, supondo constante a altura, varia na razão directa da idade, ao passo que o valor negativo de  $r_{14}$  nos revela que ela varia na razão inversa, desde que não supramos a influência que é devida aos outros factores correlacionados com a idade e em especial à altura.

Os sinais positivo e negativo, respectivamente de  $r_{14.3}$  e de  $r_{14}$ , permitem-nos ainda afirmar que a percentagem de carrasca varia na razão inversa da intensidade dos crescimentos em altura.

A mudança do sentido da variação da percentagem de casca com a idade, conforme se fixe a altura ou não, está bem evidente no gráfico 3.

A linha de regressão simples está traçada no sentido negativo, em relação à variação de percentagem de carrasca, sendo o seu coeficiente angular  $b_{14} = -0.179$ ; depois de fixarmos a altura na sua média, obtivemos uma linha de regressão parcial com um coeficiente angular  $b_{14.3} = 0.087$ , que rodou dum ângulo apreciável, sobre a primeira, passando do sentido negativo para o positivo, como o indica o seu coeficiente angular.

Resumindo, podemos então dizer:

a) que a correlação existente entre a altura e a idade,  $r_{34} = 0.855$  encobre de tal modo a verdadeira relação existente entre a idade e a



percentagem de carrasca, que a leva a mudar do sentido positivo para o sentido negativo;

b) que a percentagem de carrasca varia inversamente com a intensidade dos crescimentos em altura;

c) que ainda não temos elementos seguros para nos manifestarmos sobre a importância da idade, como factor de avaliação da percentagem de casca.

3.<sup>a</sup> — Dos caracteres ensaiados, depois da altura, é a largura da copa o factor que se nos afigura de maior importância para a avaliação de percentagem de casca. Contudo, o coeficiente de determinação  $R^2_{1.35} = 0.523$ , ainda que superior a  $R^2_{1.23}$  e  $R^2_{1.34}$ , é ainda pouco elevado, pois apenas 52,3% das variações de percentagem de carrasca se podem atribuir à altura do tronco e à largura da copa, suposta, é claro, a correlação linear.

Como  $r^2_{13} = 0.483$ , podemos afirmar que 48,30% das variações de percentagem de casca são devidas à altura e 51,70 a outros factores não considerados. Pelo facto de se considerar a largura da copa, o coeficiente de determinação aumentou em  $52,30 - 48,30 = 4,00\%$ .

A importância relativa deste acréscimo avalia-se pelo coeficiente de determinação parcial,  $r^2_{13.3} = \frac{4.00}{51.70} = 0.077$ .

Temos partido do principio de ser rectilínea a linha de regressão, mas, como a relação entre as duas variáveis mostra certa tendência curvilínea, aparentando ser parabólica a linha de regressão, impõe-se investigar se o problema terá de ser resolvido pela aplicação dos métodos baseados no principio da correlação curvilínea, ou se, pelo contrário, podem ser adoptados, sem erro apreciável, os resultados obtidos, segundo os principios da correlação rectilínea.

Com este objectivo, ajustou-se uma parábola aos valores observados, para o que se seguiu o método dos mínimos quadrados.

Nestas condições, a equação de regressão é do tipo:

$$Y = a + b X + c X^2$$

Os parâmetros  $a$ ,  $b$  e  $c$  foram determinados a partir das equações normais:

$$\begin{cases} b \sum (x^2) + c \sum (x^3) - \sum (xy) = 0 \\ b \sum (x^3) + c \sum (x^4) - \sum (x^2y) = 0 \end{cases}$$

sendo  $a = My - b Mx - c Mx^2$

Para o cálculo destes parâmetros empregámos o impresso adoptado pela Divisão de Estudos e Ordenamento (modelo 23-1937).

Os valores obtidos, depois de descodificados, foram:

$$a = 49.492$$

$$b = -4.290$$

$$c = 0.161$$

aos quais corresponde a equação de regressão:

$$X_1 = 49.492 - 4.290 X_5 + 0.161 X_5^2 \quad (7)$$

Conhecida esta equação, o índice de correlação corrigido  $\bar{p}_{15}$  é dado pela expressão:

$$\bar{p}_{15}^2 = 1 - \left( \frac{\sigma_z^2}{\sigma_1^2} \right) \left( \frac{n-1}{n-m} \right)$$

onde:  $\sigma_z$  é o d. q. m. dos resíduos  $z = X_1 - f(X_5)$ ;  $\sigma_1$ , o d. q. m. de  $X_1$ , já calculado;  $n$ , o número de observações, e  $m$ , o número de constantes da equação de regressão.

Calculou-se  $\sigma_z$  pela fórmula:

$$\sigma_z = \left( \frac{\sum (z^2) - n (Mz^2)}{n} \right)^{1/2}$$

Para determinar os valores de  $f(X_5)$ , fez-se o traçado da curva de regressão no gráfico 4.

Em cada ficha inscreveram-se os valores de  $z$  e  $z^2$ .

Feitas as somas, achámos, respectivamente:

$$\sum (z) = 0.6 \quad \text{e} \quad \sum (z^2) = 6452.8$$

Substituindo valores nas fórmulas anteriores, temos:

$$\sigma_z = \frac{6452.44}{198} = 5.709$$

e

$$\bar{p}_{15}^2 = 1 - \frac{(5.709)^2}{(7.842)^2} - \frac{198-1}{198-3} = 1 - 0.535 = 0.465$$

donde

$$\bar{p}_{15} = 0.682$$

Para calcular o valor do coeficiente de correlação, também corrigido, empregamos a fórmula:

$$\bar{r}_{15}^2 = 1 - (1 - r_{15}^2) \left( \frac{n-1}{n-2} \right)$$

Substituindo valores, obtivemos:

$$\bar{r}_{15}^2 = 0.453$$

Comparando  $\bar{p}_{15} = 0.463$  com  $\bar{r}_{15} = 0.453$ , podemos concluir que, tomando em consideração a natureza curvilínea da relação entre as duas variáveis, a variação da percentagem de casca, que pode ser explicada pela largura da copa, aumentou de 1,8% em relação à variação total.

Suposta a regressão linear, a correlação entre percentagem de casca e largura de copa é  $\bar{r}_{15} = 0.673$ ; sendo introduzido o termo do segundo grau, a correlação passa a ser  $\bar{p}_{15} = 0.682$ .

¿ Será o pequeno acréscimo verificado *significante*?

Só podemos chegar a uma conclusão segura recorrendo à *análise de variância*.

Para fazer esse estudo, calculamos as somas dos quadrados dos resíduos relativas à linha de regressão e à curva de regressão, tendo obtido os seguintes valores:

$$(1 - \bar{r}_{15}^2) (\sum x_i^2) = 0.547 \times 611.908 = 334.714$$

$$(1 - \bar{p}_{15}^2) (\sum x_i^2) = 0.535 \times 611.908 = 327.371$$

Os graus de independência para a linha de regressão e para a curva de regressão são, respectivamente:

$$198 - 2 = 196 \quad \text{e} \quad 198 - 3 = 195$$

Calculadas, no quadro C, as *variâncias* relativas aos resíduos da curva de regressão e à diferença dos resíduos das duas linhas de regressão, fez-se a sua comparação para investigar se há ou não *significância*.

O processo consiste em calcular o valor de *F* (quociente das duas *variâncias*) e compará-lo com os valores dados pela tabela de distri-

bução de  $F$  para os limites 5% e 1% (Snedcor, 1938 — Tabela 10,3) (1).

Como obtivemos  $F = 4.37$ , valor este superior a  $F_{5\%} = 3.89$ , mas inferior a  $F_{1\%} = 6.76$ , só para o nível de significância 5% é que podemos afirmar que as duas linhas de regressão são diferentes.

Devemos, pois, preferir a linha parabólica à linha recta, como de resto já tínhamos previsto: no entanto, a melhoria de ajustamento conseguida com a curva de regressão não é muito apreciável, visto que o nível de significância 1% não foi atingido.

### QUADRO C

#### Análise de variância — Ajustamento da linha de regressão parabólica

Natureza da variação	Graus de independência	Soma dos quadrados	Variância	F
Resíduos da linha de regressão	196	334.714		
Resíduos da curva de regressão	195	327.871	1.679	
Diferença	1	7.843	7.843	4.37

Analisemos agora mais pormenorizadamente o modo como os factores estudados influem na variação da percentagem de casca.

Em pinheiros com a mesma altura, a percentagem de carrasca pode considerar-se como independente do D. A. P. Esta circunstância é uma consequência da percentagem de casca aumentar com a sua espessura e diminuir com o diâmetro do lenho, e de ambas estas quantidades variarem no mesmo sentido do D. A. P. (2).

(1) Nestas tabelas entram-se com os valores de  $n_1$  e  $n_2$ , que são os graus de independência correspondentes respectivamente à maior e menor variância.

(2) Na segunda parte do nosso trabalho resolvemos considerar, em vez do D. A. P., a percentagem de espessura de casca, tomada em relação ao D. A. P. s/casca: obtivemos assim, um factor fortemente correlacionado com a variável dependente.

A percentagem de casca varia na razão inversa da altura da árvore, e, assim, todos os factores que influem no crescimento em altura e, em especial, a natureza e a profundidade do solo, teem uma influência apreciável na percentagem de casca.

O modo de tratamento também tem uma influência grande na formação da casca. Vimos já que a largura da copa é um factor de certa importância na apreciação da percentagem de casca.

O facto dos pinheiros da Mata de Foja apresentarem, em média, percentagens de casca sensivelmente mais elevadas do que as verificadas na Mata de Leiria devemos-lo attribuir, em parte, ao pinhal se encontrar, em geral, bastante mais denso na primeira mata do que na segunda.

Acérca da influência da densidade do arvoredado e da idade, na variação da percentagem de casca no pinheiro bravo, Sousa Pimentel no seu livro *Os Nossos Pinheiros* diz o seguinte:

«Nas árvores ainda novas, com 20 a 40 anos, a casca forma grande parte do volume do tronco, sobretudo quando elas crescem muito bastas, pois neste caso chegam a ter em casca 30% do volume total do tronco e nas de mais idade, com 40 a 80 anos e nas mesmas condições, 20%, mas, quando os pinheiros dispõem de bastante espaço, então apresentam menos casca, porque esta torna-se sempre mais delgada, e o mesmo succede com as árvores de mais de 80 anos, que já teem perdido a maior força de crescimento, as quais não teem mais de 10% de casca.»

Isto vem confirmar o que foi dito acérca da importância da largura da copa na avaliação da percentagem de casca.

Convém observar que os números acima transcritos, que julgo referirem-se à Mata de Leiria, se afastam muito dos valores médios observados para as idades indicadas. Um pinheiro bravo com menos de 10% de casca deve ser raro entre nós.

Para a Mata de Leiria o engenheiro silvicultor J. C. Freire Te-mudo (1) em 1445 árvores achou uma média de 31% de casca; com 10% de casca apenas encontrou duas.

Na Mata de Foja em 198 árvores, algumas das quais com 110 anos, o mínimo que encontrámos foi 18%, e a média verificada foi 39,2%. Visto em ambas as amostras predominarem as árvores novas, a diferença apreciável entre aquelas duas médias pode ser explicada, não só

(1) *Percentagem de casca no pinheiro bravo* (Relatório final do curso de Engenheiro Silvicultor, 1938).



pelas condições agrológicas, mas também pela maior densidade do arvoredo da Mata de Foja.

Ainda para a Mata de Foja, o engenheiro António Alberto da Cunha Rei (1) organizou uma tabela de percentagem de casca, segundo classes de diâmetro, que transcrevemos abaixo. Por ela pudemos verificar também os elevados valores médios de percentagem de carrasca relativos à Mata de Foja:

Classes de diâmetro	Percentagens de casca
0.10	45,2 %
0.15	40,3 %
0.20	36,7 %
0.25	33,7 %
0.30	31,2 %
0.35	29,2 %
0.40	27,5 %
0.45	26,0 %
0.50	24,7 %
0.55	23,8 %
0.60	23,2 %
0.65	23,0 %

Quanto à idade, tem ela uma importância mais restrita do que a que lhe é atribuída, se suprimirmos as variações que podem ser explicadas pela altura. O facto de se verificarem nas árvores velhas valores baixos para a percentagem de casca é devido principalmente à maior altura dessas árvores.

Em árvores de igual altura e com a mesma largura de copa verifica-se, em geral, maior percentagem de casca nas que forem mais idosas, sendo afinal este facto uma consequência de serem menores os crescimentos em altura e, portanto, de serem piores as condições agrológicas, a que estão sujeitas essas árvores.

Nos pinheiros muito velhos a casca apresenta-se unida e avermelhada, em consequência das camadas mais externas se soltarem em pequenas placas. Só quando as árvores chegam a este estado é que o factor *idade* deve influir directamente no sentido da diminuição da percentagem de carrasca, independentemente da altura das árvores.

(1) *Coefficientes de forma* (Relatório final do curso de Engenheiro Silvicultor, 1928).



### Equação de regressão múltipla (quatro variáveis)

O valor, que obtivemos para o coeficiente de determinação entre percentagem de casca, altura e largura da copa,  $R^2_{1.35} = 0.523$ , é baixo para nele basearmos a construção duma tabela de percentagem de casca.

Para avaliarmos a influência da idade, considerada conjuntamente com os factores atrás enunciados (altura do tronco e largura de copa), torna-se necessário calcular a equação de regressão para as três variáveis independentes:

$$X_1 = a_{1.345} + b_{13.45} X_3 + b_{14.35} X_4 + b_{15.34} X_5 \quad (8)$$

O processo mais rápido para se obterem os coeficientes da equação de regressão consiste em exprimir as equações normais em unidades padrões e resolver depois o sistema por elas definido.

No caso presente, em que as variáveis são quatro, temos três equações normais:

$$\begin{aligned} r_{13} &= \beta_{13} + \beta_{14} \cdot r_{34} + \beta_{15} \cdot r_{25} \\ r_{14} &= \beta_{13} \cdot r_{34} + \beta_{14} + \beta_{15} \cdot r_{45} \\ r_{15} &= \beta_{13} \cdot r_{35} + \beta_{14} \cdot r_{45} + \beta_{15} \end{aligned}$$

Resolveu-se este sistema pelo método de Doolittle.

Os coeficientes da equação (8) calcularam-se pela fórmula geral:

$$b_{12.3\dots n} = \left( \frac{\sigma_1}{\sigma_2} \right) \beta_{12.3\dots n} \quad (9)$$

e o termo independente pela expressão analítica:

$$a = M_{X_1} + b_{13} M_{X_3} + b_{14} M_{X_4} + b_{15} M_{X_5}$$

Os valores obtidos, ainda codificados, foram os seguintes:

$$b_{13.45} = 1.133 \times -0.713 = -0.808$$

$$b_{14.35} = 1.137 \times 0.422 = 0.480$$

$$b_{15.34} = 1.487 \times -0.436 = -0.648$$

$$e \quad a = 6.748 + 3.583 - 1.973 + 1.813 = 10.171$$

A equação de regressão codificada é então:

$$X_1 = 10.171 - 0.808 X_3 + 0.480 X_4 - 0.648 X_5$$

Determinada a equação de regressão, calcula-se o coeficiente de múltipla determinação pela fórmula (5):

$$R_{1.345}^2 = \frac{766.1885}{1352.672} = 0.566$$

Para avaliar a sucessiva melhoria do ajustamento conseguido pela introdução de cada uma das novas variáveis independentes, calcularam-se os resíduos quadráticos médios para as três equações de regressão.

Pela aplicação da fórmula geral

$$S_{1.234 \dots n}^2 = \sigma_1^2 (1 - R_{1.234 \dots n}^2) \left( \frac{n-1}{n-m} \right) \quad (10)$$

determinaram-se os valores que foram inscritos no quadro D, onde veem também indicados os coeficientes de correlação múltipla e de determinação total.

QUADRO D

Variável dependente	Variáveis independentes	Resíduo quadrático médio	Coef. de múltipla correlação	Coef. de determinação total
$X_1$	$X_3$	5.666	—	—
$X_1$	$X_3, X_5$	5.442	0.723	0.523
$X_1$	$X_3, X_5, X_4$	5.190	0.752	0.566

$X_4$  = Percentagem de casca.

$X_3$  = D. A. P.

$X_4$  = idade.

$X_5$  = largura da copa.

Como é evidente, por cada variável introduzida diminuiu o resíduo quadrático médio e aumentou o coeficiente de correlação.

### Coefficientes de determinação parcial (quatro variáveis)

Para avaliar a importância individual de cada variável, tomando simultaneamente em consideração o efeito dos outros factores, calcularam-se os coeficientes de determinação parcial, para o que se aplicou a fórmula geral:

$$r_{12,34 \dots n}^2 = 1 - \frac{(1 - R_{1,234 \dots n}^2)}{(1 - R_{1,34 \dots n}^2)} \quad (11)$$

Obtiveram-se assim os seguintes valores:

$$r_{13,45}^2 = 1 - \frac{1 - 0.566}{1 - 0.448} = 0.214$$

$$r_{14,35}^2 = 1 - \frac{1 - 0.566}{1 - 0.523} = 0.089$$

$$r_{15,34}^2 = 1 - \frac{1 - 0.566}{1 - 0.502} = 0.129$$

Podemos concluir, em face destes números, que a altura do tronco é o factor mais influente, pois a ele se podem atribuir 21,4 % das variações da percentagem de casca deixadas de explicar pela largura de copa e idade. Seguem-se, por ordem de importância, a largura da copa com 12,9 % e a idade com 8,9 %. (Vide quadro E).

QUADRO E

Factores já considerados	Factor acrescentado	Coef. de correlação parcial	Coef. de determinação parcial
X <sub>3</sub> , X <sub>4</sub>	X <sub>5</sub>	0.359	0.129
X <sub>2</sub> , X <sub>5</sub>	X <sub>4</sub>	0.299	0.089
X <sub>4</sub> , X <sub>5</sub>	X <sub>3</sub>	0.463	0.214

## Conclusão

Este estudo prévio, feito com o principal objectivo de escolher os factores de maior influência na percentagem de casca, para nêles basear a organização duma tabela, não nos permitiu tirar uma conclusão definitiva.

O valor do coeficiente de determinação total,  $R^2_{1.345} = 0.566$  é evidentemente baixo, pois ficam por explicar 43,4 % das causas de variação da percentagem de casca.

¿ Como poderemos aumentar o rigor da avaliação?

Há dois caminhos a seguir:

1.º — Melhorar o ajustamento, modificando a superfície plana numa superfície curva.

2.º — Procurar encontrar outros factores que se mostrem mais influentes.

Ainda que o primeiro caminho tenha certo fundamento no nosso caso, visto termos verificado ser curvilínea a relação entre a largura da copa e a percentagem de casca, ele não nos conduziria a resultados satisfatórios. Se tomarmos em consideração a natureza curvilínea da equação de regressão, diminuiríamos sem dúvida o residuo quadrático médio, mas nunca conseguiremos um rigor que se aproxime do que pretendemos obter.

Torna-se, por isso, necessário introduzir um factor de maior influência, e, depois, se as circunstâncias o aconselharem, modificar a superfície plana numa superfície curva, por qualquer dos métodos da múltipla regressão curvilínea.

É o que vamos procurar conseguir na segunda parte do nosso trabalho, com o fim de podermos organizar uma tabela de percentagem de casca suficientemente rigorosa.

## II PARTE

### Organização duma tabela de percentagem de casca

#### Generalidades

O antagonismo, que encontrámos entre as várias condições a que deve satisfazer uma boa tabela de percentagem de casca, tornou particularmente difícil a sua elaboração. Há três pontos fundamentais a considerar:

1.º — As variáveis independentes a utilizar nunca devem representar menos de 80% das causas de variação dos valores da variável dependente.

2.º — As variáveis independentes devem ser, tanto quanto possível, fáceis de medir directamente.

3.º — O número de variáveis independentes a empregar deve ser o menor possível, para tornar a tabela de fácil aplicação.

Quando existem um ou dois caracteres fáceis de medir directamente e fortemente correlacionados com o factor que se pretende avaliar, o problema apresenta-se simplificado por natureza.

Infelizmente, para a percentagem de casca, as coisas passam-se diversamente; nenhum dos caracteres morfológicos da árvore se apresenta *fortemente significativa*.

São muitos os factores que interveem e, em virtude da correlação existente entre eles, acontece que o coeficiente de determinação múltipla, pouca aumenta em relação a cada factor adicionado.

Estão, portanto, em absoluto antagonismo os três quesitos enunciados. Tornou-se, por isso, necessário encontrar uma solução intermédia dentro das conclusões a que nos conduziu o nosso trabalho de investigação.

Na solução por nós adoptada tomou-se como base o primeiro quesito e procurou-se adaptar-lhe, tanto quanto possível, o segundo e o terceiro.

Este modo de resolver o problema apresenta, sem dúvida, inconveniente de ordem prática, como adiante se verá, mas, em face das conclusões a que chegámos, pareceu-nos o melhor.

### Escolha dos factores a considerar. Análise de variância

Vimos, pelas investigações anteriores, que a variação da percentagem de casca com o D. A. P., para cada classe de alturas era praticamente nula; concluímos também que a altura do tronco, a idade e a largura da copa, só por si, eram insuficientes para a organização duma tabela de percentagem de casca. Esta circunstância obrigou-nos a prosseguir as investigações, que, por sua vez, nos levaram a considerar uma nova variável independente: a espessura de casca expressa em percentagem do D. A. P. s/ casca, variável esta que designamos por  $X_6$ .

Para avaliar a importância deste factor, tivemos de calcular os coeficientes de simples correlação de  $X_6$  com  $X_1$ ,  $X_3$ ,  $X_4$  e  $X_5$ , para o que applicámos a fórmula (1).

#### QUADRO F

##### Coefficientes de correlação simples

	1	3	4	5
6	0.839	-0.487	-0.425	-0.588
5	-0.670	0.787	0.774	
4	-0.525	0.855		
3	-0.695			

1 = casca (%).

3 = altura.

4 = idade.

5 = largura da copa.

6 = espessura da casca (%).

No quadro F inscrevemos os coeficientes de simples correlação entre as cinco variáveis que estamos considerando.

O alto valor de  $r_{16}$  faz-nos prever resultados animadores, quanto à importância do novo factor introduzido, na avaliação da percentagem da casca.

Para investigar a natureza rectilínea ou curvilínea da relação entre  $X_1$  e  $X_6$ , calculámos a linha de regressão:

$$X_1 = 7.430 + 2.211 X_6$$



e traçamo-la no gráfico 5, juntamente com a linha das médias. O exame deste gráfico permite-nos afirmar que a linha recta é o melhor ajustamento.

A *análise de variância*, de que falaremos mais adiante, conduziu-nos a conclusão análoga (Vide quadro K).

Resta agora verificar se o facto de entrarmos com a nova variável  $X_6$  não vem tornar desnecessário considerar alguns dos factores já empregados. Com este objectivo calcularam-se os coeficientes de múltipla determinação, para o que aplicámos mais uma vez a fórmula geral (5).

Para isso tivemos de determinar os respectivos coeficientes de regressão, seguindo-se o método de Doolittle.

Aplicámos também este método para a resolução das equações normais para 5 variáveis, cujas incógnitas são os coeficientes de regressão, necessários ao cálculo de  $R^2_{1.3456}$ .

Os valores obtidos foram os seguintes:

$$R^2_{1.345} = \frac{766.1885}{1352.672} = 0.5664 \quad (a)$$

$$R^2_{1.346} = \frac{1126.7411}{1352.672} = 0.8329 \quad (b)$$

$$R^2_{1.350} = \frac{1098.5823}{1352.672} = 0.8121 \quad (c)$$

$$R^2_{1.456} = \frac{1016.682}{1352.672} = 0.7509 \quad (d)$$

$$R^2_{1.3456} = \frac{1128.8282}{1352.672} = 0.8351 \quad (e)$$

Para escolher os factores que devem ser considerados no cálculo de  $X_1$ , vamos investigar, por meio da análise de variância, se é significativo o acréscimo sofrido por qualquer dos coeficientes de correlação múltipla:  $R_{1.345}$ ,  $R_{1.346}$ ,  $R_{1.350}$  e  $R_{1.456}$ , pela introdução da quarta variável independente que falta considerar em cada um deles.

Como (b) é o valor que mais se aproxima de (e), começámos por investigar se  $R_{1.346}$  e  $R_{1.3456}$  são *significativamente diferentes*.

Com este fim, calculámos  $(1 - R^2_{1.346})$  o  $(1 - R^2_{1.3456})$ , valores estes proporcionais às somas dos quadrados dos resíduos das respectivas equações de regressão, aos quais correspondem os graus de independência:

$$N - p - 1 = 194 \quad \text{e} \quad N - p' - 1 = 193$$

representando  $N$  o número de observações e  $p$  e  $p'$  o número de variáveis independentes de cada uma das equações.

Comparando o valor de  $F$  fornecido pela *análise de variância* (Vide quadro G) com os valores  $F_5^0 = 3.89$  e  $F_4^0 = 6.77$ , verificámos que  $R_{1,356}$  e  $R_{1,3456}$  não são *significativamente diferentes* e, portanto, que não vale a pena considerar o factor  $X_5$  (largura da copa).

Seguidamente vamos ensaiar  $R_{1,356}$ , por ser o que sucede em ordem de grandeza.

Calculámos do mesmo modo a soma dos quadrados e os graus de independência, que foram também inscritos no quadro G. O valor de  $F$ , que obtivemos, sendo muito superior a  $F_1^0 = 6.77$ , indica-nos que o acréscimo sofrido por  $R_{1,346}$  pela introdução do factor  $X_1$  (idade) é *fortemente significativa*.

Como os graus de independência são os mesmos, podemos afirmar com mais forte razão, sem recorrer à análise de variância, que  $R_{1,456}$  e  $R_{1,345}$  são *significativamente diferentes* de  $R_{1,3456}$ .

Devemos pois considerar os factores  $X_6$ ,  $X_3$  e  $X_1$ , mas desprezar o factor  $X_5$ .

#### QUADRO G

**Análise de variância — Significância do acréscimo sofrido pelo coeficiente de múltipla correlação pela introdução dum novo factor**

Natureza da variação	Soma dos quadrados	Graus de independência	Variância	F
Resíduos: equação de regressão f ( $X_1, X_3, X_4, X_6$ )	0.1671	194		
Resíduos: equação de regressão f ( $X_1, X_3, X_4, X_5, X_6$ )	0.1649	198	0.00085	
Diferença	0.0022	1	0.0022	2.6
Resíduos: equação de regressão f ( $X_1, X_3, X_5, X_6$ )	0.1879	194		
Resíduos: equação de regressão f ( $X_1, X_3, X_4, X_5, X_6$ )	0.1649	198	0.00085	
Diferença	0.0230	1	0.0230	27

GRÁFICO 5

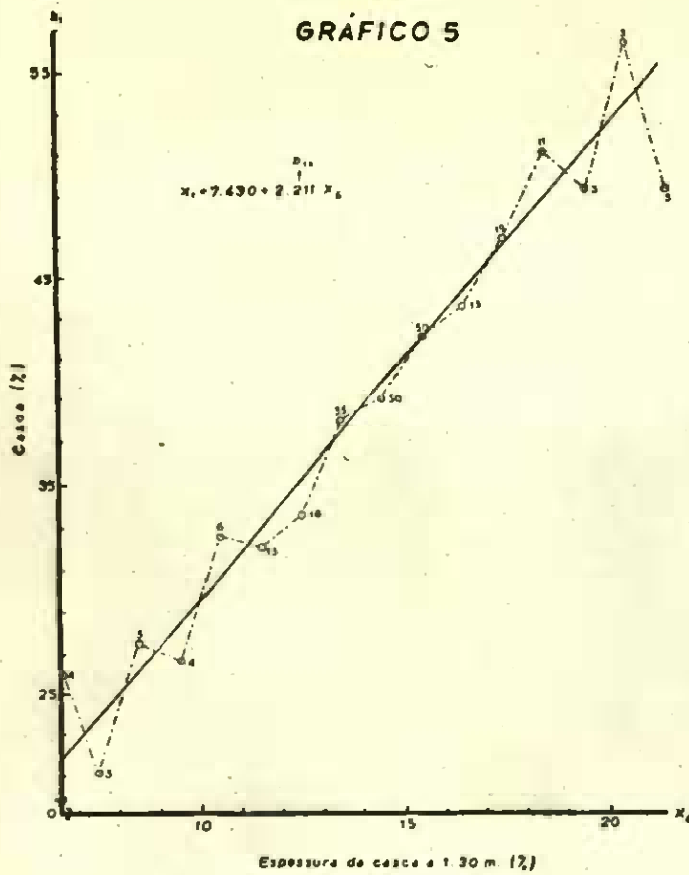
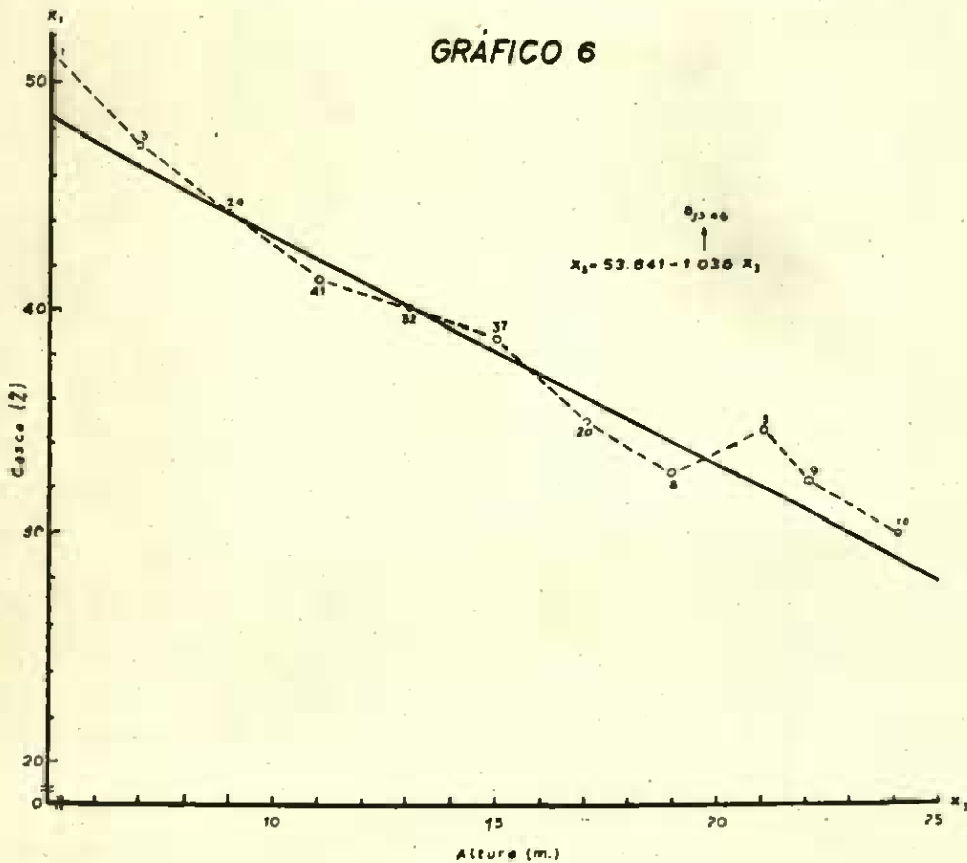
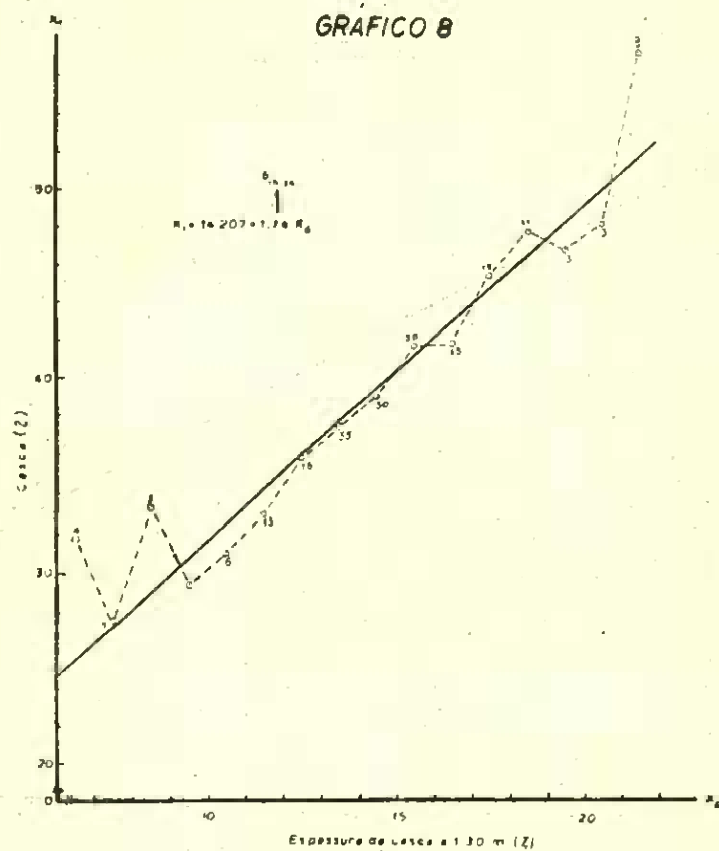
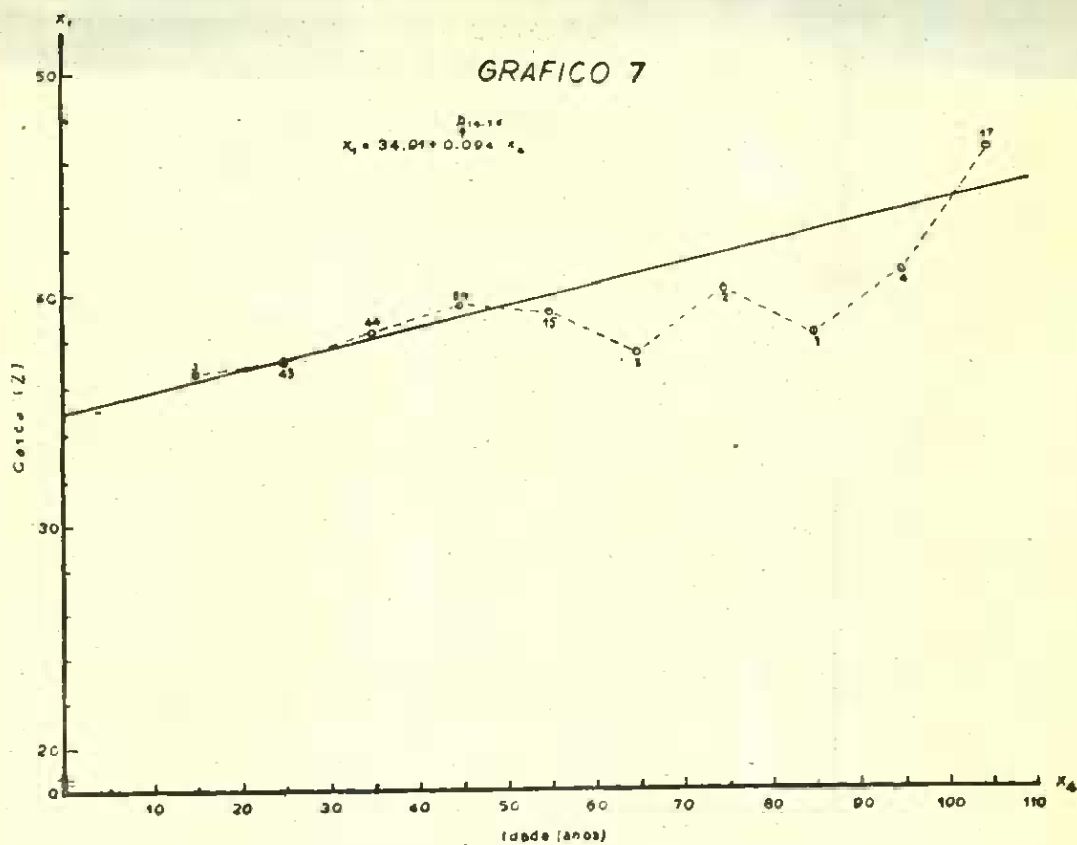


GRÁFICO 6





O resultado, a que chegámos, não significa de modo algum que a percentagem de casca seja independente da largura da copa; apenas nos leva a concluir que as causas de variação da percentagem de casca, devidas a este factor, são também explicadas pela percentagem de espessura de casca ( $X_6$ ) e, portanto, que, uma vez considerado o factor  $X_6$ , será desnecessário considerar o factor  $X_5$ .

### Investigação da natureza linear ou curvilínea da equação de regressão — Análise gráfica e análise numérica

Escolhidas as variáveis independentes, resta-nos agora averiguar se será lícito basear-nos nos princípios da correlação linear para a determinação da equação de regressão.

Supondo rectilíneas as relações entre cada uma das variáveis independentes e a variável dependente, obtivemos a seguinte equação de regressão:

$$X_1 = 24.24 + 1.74 X_6 - 1.036 X_3 + 0.094 X_4 \quad (12)$$

Para investigarmos se se torna necessário recorrer a uma equação de regressão curvilínea ou se satisfaz a equação de regressão proposta, vamos determinar as linhas das médias dos resíduos.

Com o fim de calcularmos os valores de  $X_1$ , dados pela equação (12), construímos um nomograma, que vem reproduzido na página seguinte.

Os resíduos  $z$  (diferenças entre os valores observados e os valores calculados) foram inscritos nas fichas e distribuídos sucessivamente por classes de  $X_3$ ,  $X_4$  e  $X_6$ .

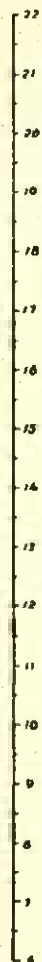
Nos gráficos 6, 7 e 8 marcaram-se em ordenadas e a partir das linhas de regressão parcial as médias dos resíduos, calculadas nos quadros anteriores.

Traçadas as linhas das médias dos resíduos, não observámos nenhuma tendência acentuadamente curvilínea que justificasse a necessidade de modificar a superfície plana numa superfície curva.

No entanto, para termos tôdas as garantias dum bom ajustamento e fugirmos ao critério pessoal a que o exame dos gráficos poderiam conduzir, vamos recorrer à *análise de variância* e investigar se não é linear a relação entre a variável dependente e qualquer das variáveis independentes consideradas (*Fisher-Test for non-linearity*).

# NOMOGRAMA

ESPESSURA DA CASCA  
(%)



IDADE  
(anos)

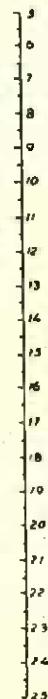


CASCA  
(%)



$$X_1 = 24.24 + 1.74 X_2 - 1.04 X_3 - 0.09 X_4$$

ALTURA  
(m)





O método consiste em dividir primeiramente a soma total dos quadrados dos desvios entre as  $m$  classes e outra relativa aos desvios das  $n_s$  observações dentro de cada classe, isto é:

$$\Sigma (Y - \bar{Y})^2 = \Sigma_1^m \Sigma' (Y - \bar{Y}_s)^2 + \Sigma_1^m n_s (\bar{Y}_s - \bar{Y})^2$$

em que,  $\bar{Y}$  representa a média geral,  $\bar{Y}_s$  a média em cada classe e o símbolo  $\Sigma_1^m \Sigma'$  a soma total  $\Sigma$ .

A equação dos correspondentes graus de independência é:

$$(N-1) = (m-1) + [(N-1) - (m-1)]$$

onde  $N$  representa o número total de observações.

Por sua vez, a «soma dos quadrados» relativa aos desvios entre as classes, pode desdobrar-se em duas partes: uma, correspondente à soma dos quadrados dos desvios dados pela linha de regressão, e outra relativa aos desvios das médias das classes para a linha de regressão (resíduos), conforme mostra o quadro H, onde veem também inscritos os graus de independência correspondentes.

No mesmo quadro vem indicado o modo como os diversos valores foram calculados.

#### QUADRO H

##### Disposição do cálculo para a análise de variância

Natureza da variação		Soma dos quadrados	Graus de indep.
Linha de regressão. .	Entre as classes	$r^2 \Sigma (Y - \bar{Y})^2$ $\Sigma_1^m n_s (\bar{Y}_s - \bar{Y})^2 =$ $= r^2 \Sigma (Y - \bar{Y})^2$	1. . . .
Resíduos. .	Diferença .		m - 1
Resíduos dentro das classes . . . . .		$\Sigma_1^m \Sigma' (Y - \bar{Y}_s)^2 = (1 - r^2) \Sigma (Y - \bar{Y})^2$	m - 2 .
Total . . . . .		$\Sigma (Y - \bar{Y})^2$	N - 1

Para maior simplicidade do cálculo, exprimimos tôdas as somas dos quadrados em função do *total* (valor este já calculado), para o que recorreremos à *razão de correlação* ( $r_{yx}$ ) entre a variável dependente e cada uma das variáveis independentes. Para a sua determinação aplicamos a fórmula:

$$1 - r_{yx}^2 = \frac{\sum (Y - \bar{Y}_s)^2}{\sum (Y - \bar{Y})^2} \quad (14)$$

e, para o cálculo do numerador da fracção, utilizámos a expressão analítica:

$$\sum (Y - \bar{Y}_s)^2 = \sum (Y^2) - \sum \frac{T_s^2}{n_s} \quad (15)$$

onde  $T_s^2$  representa o quadrado da soma dos  $n_s$  valores de  $Y$  em cada uma das  $m$  classes de  $X$ .

Para apreciar a significância dos desvios para a linha recta compara-se a variância dos resíduos das médias das classes para a linha de regressão com os resíduos internos das classes.

Quanto maior fôr o primeiro em relação ao segundo, maior será o desvio para a linha recta. Averigua-se se este desvio é ou não significativo, comparando o valor de  $F$  relativo ao quociente das duas variâncias com o valor de  $F_{5\%}$  dado pelas tabelas.

### QUADRO I

#### Análise de variância — Ajustamento da equação de regressão

$$f(X_1, X_3)$$

Natureza da variação		Soma dos quadrados	Graus de indep.	Variância
Linha de regressão . .	Entre as classes	653.34	1	653.34
		712.86	10	
Resíduos. .		59.52	9	6.61
Resíduos dentro das classes . . . . .		639.81	187	3.42
Total . . . . .		1852.67	197	—

$$F = \frac{6.61}{3.42} = 1.93$$

$$F_{5\%} = 1.93$$

Da análise de variância do quadro I concluímos que podemos considerar rectilínea a relação entre  $X_1$  e  $X_3$ , visto que o nível de significância 5 % foi atingido, mas não ultrapassado.

QUADRO J

Análise de variância — Ajustamento da equação de regressão  
f(X<sub>1</sub>, X<sub>4</sub>)

Natureza da variação	Soma dos quadrados	Graus de indep.	Variância
Linha de regressão . .	373.34	1	373.34
Entre as classes	457.20	9	
Resíduos. .	83.86	8	10.48
Resíduos dentro das classes . . . . .	895.47	188	4.76
Total . . . . .	1352.67	197	—

$$F = \frac{10.48}{4.76} = 2.22$$

$$F_{5\%} = 1.99$$

$$F_{1\%} = 2.61$$

A análise de variância do quadro J leva-nos a concluir que é curvilínea a relação entre X<sub>1</sub> e X<sub>4</sub>.

No entanto, como o valor que obtivemos para F está mais próximo de F<sub>5%</sub> do que de F<sub>1%</sub> e como a idade é de todos os factores considerados o que menos influe na percentagem de casca, resolvemos não atender à natureza curvilínea da linha de regressão.

QUADRO K

Análise de variância — Ajustamento da equação de regressão  
f(X<sub>1</sub>, X<sub>6</sub>)

Natureza da variação	Soma dos quadrados	Graus de indep.	Variância
Linha de regressão . .	952.28	1	952.28
Entre as classes	968.51	15	
Resíduos. .	16.23	14	1.16
Resíduos dentro das classes . . . . .	384.16	182	2.11
Total . . . . .	1352.67	197	—

$$F = \frac{2.11}{1.16} = 1.82$$

$$F_{5\%} = 2.17$$

Do exame do quadro K concluímos que, para a relação entre  $X_1$  e  $X_8$ , é muito bom o ajustamento da linha recta (Vide gráfico 5).

O estudo, que acabamos de fazer, vem justificar as conclusões que tirámos da análise dos gráficos 6, 7 e 8.

Não vemos, pois, necessidade de recorrer a aproximações, pelo que adoptámos a equação de regressão proposta.

### Considerações finais

Como, no nosso caso, em que os eixos são cinco, as leituras feitas no nomograma são morosas, construímos uma tabela que vem indicada na página seguinte.

Visto nem sempre termos o conhecimento exacto das árvores, para tornar a tabela de mais fácil aplicação, considerámos apenas quatro classes de idade. Para cada classe organizámos uma tabela de percentagem de casca de duas entradas, marcando horizontalmente os valores de percentagem da espessura de casca e verticalmente as alturas.

A idade, em relação a estes dois factores, é um carácter de muito menor importância: por isso não são apreciáveis os erros que se cometem, ao considerarem-se apenas quatro classes de idade.

Com esta tabela podem-se calcular as percentagens de carrasca em árvores de qualquer idade, sendo os resultados manifestamente seguros. Com efeito, conseguimos um coeficiente de múltipla correlação  $R_{1.346} = 0.913$ , a que corresponde um coeficiente de determinação total  $R^2_{1.346} = 0.833$ , resultado este apreciável, atendendo aos numerosos factores que interveem.

Permite ainda a tabela avaliar a percentagem de casca com as árvores em pé, o que tem uma enorme vantagem prática.

Com efeito, conhecendo-se pelo plano de ordenamento a idade das árvores, pode-se avaliar com mais rigor a percentagem de carrasca pelo emprêgo do nomograma, mas, se ela não é conhecida, fácil é classificá-la dentro duma das quatro classes de idade por nós consideradas para a construção da tabela.

A altura determina-se facilmente com as árvores em pé com um dendrómetro.

Para calcular as percentagens de espessura de carrasca, torna-se necessário conhecer dois elementos: o D. A. P. s/casca e a espessura da casca.

Na Estação de Experimentação do Pinheiro Bravo é já muito cor-

Tabela para a avaliação da percentagem de casca

Classes de idade (anos)	Altura (m)	Espessura de casca (% referida ao D. A. P. sobre casca)																		
		7	8	9	10	11	12	13	14	15	16	17	18	19	20	21				
21-40	6			—	38.6	40.2	42.2	43.8	45.7	47.3	49.3	51.0	52.8	54.7	56.4	58.8				
	7			—	37.5	39.0	41.0	42.7	44.6	46.2	48.1	49.8	51.7	53.5	55.2	57.1				
	8			—	36.3	37.8	39.8	41.6	43.4	45.1	46.9	48.6	50.5	52.3	54.0	56.0				
	9			—	35.0	36.7	38.7	40.4	42.2	44.0	45.7	47.4	49.4	51.1	52.9	54.8				
	10			—	33.8	35.5	37.5	39.2	41.0	42.8	44.6	46.3	48.2	49.9	51.8	53.6				
	11		30.7	32.5	34.3	36.3	38.0	39.8	41.6	43.5	45.1	47.1	48.7	50.7	52.4	54.8				
	12		29.5	31.3	33.1	35.2	36.8	38.5	40.4	42.3	44.0	46.0	47.6	49.5	51.2	53.6				
	13		28.5	30.1	31.9	34.0	36.7	37.5	39.2	41.1	42.8	44.8	46.4	—	—	—				
	14		27.1	29.0	30.7	32.7	34.5	36.4	38.0	39.9	41.7	43.7	45.2	—	—	—				
	15		26.0	27.8	29.6	31.5	33.3	35.2	36.9	38.7	40.5	42.5	44.1	—	—	—				
41-60	16		24.9	26.7	28.4	30.3	32.1	33.9	35.8	37.5	39.3	41.3	43.0	—	—	—				
	17		23.7	25.5	27.2	29.0	30.9	32.7	34.6	36.4	38.2	40.1	41.8	—	—	—				
	9		—	37.5	39.2	41.2	42.9	41.6	46.3	48.2	49.9	51.9	53.6	—	—	—				
	10		—	36.3	38.0	40.0	41.7	43.4	45.1	47.0	48.7	50.7	52.4	—	—	—				
	11		33.2	35.1	36.8	38.8	40.5	42.3	44.0	45.8	47.5	49.6	51.2	—	—	—				
	12		32.0	33.9	35.7	37.6	39.3	41.2	42.9	44.7	46.3	48.4	50.0	—	—	—				
	13		30.8	32.7	34.5	36.5	38.1	40.0	41.7	43.6	45.2	47.2	48.9	—	—	—				
	14		29.7	31.5	33.3	35.3	37.0	38.8	40.5	42.4	44.1	46.0	47.7	—	—	—				
	15		28.5	30.2	32.1	34.1	35.8	37.6	39.3	41.2	43.0	44.8	46.5	—	—	—				
	16		27.3	29.0	30.9	32.8	34.7	36.4	38.1	40.0	41.9	43.7	45.4	—	—	—				
61-80	17		26.2	27.8	29.7	31.5	33.5	35.2	37.0	38.8	40.7	42.5	44.2	—	—	—				
	18		25.0	26.7	28.5	30.3	32.3	34.0	35.9	37.7	39.6	—	—	—	—	—				
	19		23.8	25.6	27.3	29.2	31.0	32.8	34.7	36.6	38.4	—	—	—	—	—				
	20		22.7	24.5	26.2	28.0	29.7	31.6	33.4	35.4	37.2	—	—	—	—	—				
	21		21.5	23.3	25.0	26.8	28.5	30.4	32.2	34.2	36.1	—	—	—	—	—				
	22		20.3	22.2	23.9	25.6	27.3	29.2	31.0	33.0	35.0	—	—	—	—	—				
	15		30.9	32.7	34.6	36.7	38.2	40.0	41.9	43.6	45.4	—	—	—	—	—				
	16		29.7	31.5	33.4	35.4	37.0	38.8	40.7	42.5	44.9	—	—	—	—	—				
	17		28.5	30.3	32.2	34.1	35.9	37.6	39.5	41.3	43.2	—	—	—	—	—				
	18		27.3	29.1	31.0	32.9	34.7	36.4	38.3	40.1	42.0	—	—	—	—	—				
81 e mais	19		26.1	28.0	29.7	31.7	33.5	35.2	37.1	38.9	40.8	—	—	—	—	—				
	20		25.1	26.8	28.5	30.5	32.2	34.0	36.0	37.7	39.6	—	—	—	—	—				
	21		24.0	25.7	27.3	29.2	31.0	32.8	34.8	36.6	38.4	—	—	—	—	—				
	22		22.8	24.6	26.2	28.0	29.8	31.6	33.7	35.4	37.2	—	—	—	—	—				
	23		21.7	23.4	25.1	26.9	28.6	30.5	32.5	—	—	—	—	—	—	—				
	20	—	26.7	28.6	30.5	32.4	34.2	36.2	37.8	39.6	41.5	43.4	—	—	—	—				
	21	—	25.6	27.4	29.3	31.1	33.0	35.0	36.6	38.4	40.3	42.2	—	—	—	—				
	22	—	24.6	26.3	28.2	29.8	31.8	33.8	35.5	37.2	39.1	41.0	—	—	—	—				
	23	21.8	23.5	25.2	26.9	28.6	30.6	32.8	34.4	36.0	—	—	—	—	—	—				
	24	20.5	22.5	24.0	25.8	27.5	29.4	31.3	—	—	—	—	—	—	—	—				
25	19.3	21.5	23.0	24.7	26.4	28.2	30.0	—	—	—	—	—	—	—	—					
26	18.0	20.0	21.6	23.4	25.2	27.0	28.8	—	—	—	—	—	—	—	—					

rente medirem-se directamente as espessuras de casca por meio de sondagens, o que evita abater as árvores.

Fazem-se simultaneamente as medições cruzadas do D. A. P. e da espessura de casca.

Por meio duma régua de cálculo avalia-se expeditamente a percentagem da espessura de casca, em relação ao D. A. P. s/casca, mediante a operação:

$$\frac{E}{D} \times 100 \quad (16)$$

Este elemento é, sem dúvida, pouco útil, debaixo do ponto de vista prático, pois, para o conhecer, torna-se necessário efectuar duas medições e uma operação; no entanto, tivemos de utilizá-lo, devido a ser ele o que se apresenta mais fortemente correlacionado com a variável dependente.

Pode-se também construir uma tabela que nos dê percentagens de espessura de casca em função do D. A. P. e da espessura da casca, consoante a fórmula (16), o que torna muito mais expedito o cálculo daquele factor.



Quadro I

Número do registo	Casca (%) $X_1$	D. A. P. (cm) $X_2$	Altura (m) $X_3$	Idade (anos) $X_4$	Largura de copa (m) $X_5$	Espessura de casca (%) $X_6$
1	42.9	24.4	15.1	45	3.10	17.0
2	45.9	22.2	13.9	45	2.90	16.7
3	46.4	18.6	16.0	45	1.70	19.3
4	52.9	16.6	12.2	45	1.40	19.9
5	45.4	12.4	10.8	45	1.20	16.9
6	51.0	15.8	11.5	45	1.50	17.7
7	43.8	19.3	13.4	45	2.20	15.3
8	48.4	23.8	13.8	45	2.60	17.6
9	42.0	18.6	14.2	46	2.10	15.3
10	38.1	18.2	14.4	45	2.10	15.9
11	52.6	18.0	12.2	45	1.70	21.1
12	18.0	36.0	24.1	95	4.40	9.7
13	39.0	45.0	22.8	102	5.20	15.0
14	26.1	50.0	25.1	105	9.20	6.9
15	34.7	45.0	24.5	105	6.50	13.3
16	25.2	34.6	23.6	92	6.20	12.4
17	26.6	41.0	23.9	109	8.10	6.7
18	28.0	43.5	23.3	102	7.10	6.0
19	24.9	45.0	24.7	101	7.80	9.4
20	28.9	47.5	25.6	100	4.60	6.8
21	20.4	41.5	24.6	103	6.80	7.1
22	20.5	41.2	25.0	102	5.30	7.5
23	28.9	44.0	23.5	103	6.70	8.9
24	43.1	45.2	22.5	102	6.20	17.7
25	31.1	49.0	21.6	104	8.00	8.2
26	29.9	40.0	22.5	109	5.80	9.3
27	25.3	60.0	25.5	105	7.20	8.5
28	23.6	55.0	25.6	105	9.50	13.1
29	31.0	39.0	22.9	103	6.00	11.7
30	28.8	39.8	22.9	92	6.20	12.4
31	22.2	41.0	24.6	103	6.00	7.7
32	29.5	39.0	21.6	50	6.10	8.0
33	39.4	18.0	14.4	50	2.00	17.2
34	41.8	24.8	15.8	58	3.20	16.1
35	44.5	23.8	14.9	51	3.00	16.2
36	47.6	15.0	12.8	49	1.50	18.0
37	41.3	11.0	11.7	41	1.20	13.2
38	34.6	11.6	10.5	47	2.40	11.6
39	41.5	20.6	12.7	45	2.90	15.3
40	37.9	23.8	15.5	49	2.70	13.2
41	41.3	23.8	16.9	51	2.50	14.7
42	25.6	39.2	19.4	69	7.00	11.5
43	44.7	15.8	11.4	43	2.70	14.6
44	40.0	23.8	14.3	50	4.00	15.5
45	46.5	31.4	14.2	48	4.00	17.8
46	40.9	20.5	15.0	45	2.60	15.4
47	40.0	22.2	13.4	40	2.60	14.8
48	50.6	15.0	11.3	41	1.30	18.0
49	33.5	30.2	17.0	53	3.80	12.6
50	40.0	17.6	13.1	37	2.50	13.6
51	50.0	14.5	12.7	38	1.40	19.0
52	37.1	23.8	13.4	43	2.10	13.9
53	33.5	17.5	17.8	33	3.00	13.7
54	42.2	17.3	12.4	33	2.00	15.3
55	41.1	16.8	11.1	33	2.20	14.8
56	34.4	18.2	13.5	34	2.60	13.2
57	40.0	21.2	13.2	35	2.60	15.1
58	48.7	10.8	9.0	30	2.00	16.7
59	41.6	20.5	12.8	31	2.50	15.4
60	41.6	22.0	13.2	29	2.30	15.9
61	35.6	17.8	12.5	25	2.20	13.7
62	40.1	19.5	11.5	30	1.50	13.6
63	42.3	16.8	11.7	29	1.50	14.9
64	41.0	16.9	10.6	29	2.70	14.5
65	33.7	14.5	10.7	28	1.50	10.0
66	36.7	14.2	11.2	31	3.00	16.2

Quadro I (continuação)

Número do registo	Casca (%) $X_1$	D. A. P. (cm) $X_2$	Altura (m) $X_3$	Idade (anos) $X_4$	Largura de copa (m) $X_5$	Espessura de casca (%) $X_6$
67	46.1	8.7	8.5	28	1.20	15.5
68	50.0	14.9	9.1	29	1.80	24.5
69	37.5	12.4	9.2	29	1.70	12.1
70	41.2	15.5	10.5	31	2.20	14.5
71	37.0	16.8	10.7	29	2.60	13.7
72	40.6	13.6	8.5	31	2.20	11.8
73	50.5	11.5	8.8	28	1.40	17.0
74	49.2	17.2	10.2	31	2.50	17.4
75	40.5	9.6	8.8	29	1.60	13.5
76	48.8	14.5	10.5	28	1.80	15.5
77	39.8	17.5	10.6	30	2.70	13.6
78	38.3	28.2	14.5	42	4.10	14.5
79	37.5	16.0	10.0	27	2.50	11.2
80	36.6	16.9	9.5	24	2.40	11.8
81	43.6	20.1	14.2	44	2.60	17.3
82	41.2	14.6	11.6	40	17.0	12.3
83	48.0	15.8	12.2	39	17.0	16.8
84	32.4	23.0	16.0	44	3.00	14.1
85	29.7	28.0	19.6	50	2.50	14.4
86	30.6	30.2	19.7	50	3.80	14.4
87	34.9	20.2	15.9	50	2.40	14.8
88	29.6	20.9	16.1	52	2.30	12.7
89	52.0	12.6	8.9	41	1.40	17.5
90	36.1	15.5	17.1	40	1.90	15.2
91	30.6	21.0	16.8	52	3.40	14.3
92	37.8	19.3	13.8	47	2.30	16.3
93	40.0	16.7	13.4	34	2.10	15.9
94	34.2	31.2	16.6	52	3.80	15.7
95	47.5	18.2	14.6	50	1.40	18.7
96	44.0	22.0	14.4	48	2.30	17.3
97	39.0	28.0	15.9	49	2.80	13.2
98	45.0	18.5	13.8	47	1.80	17.6
99	37.7	21.4	15.0	48	3.00	15.4
100	43.4	22.8	14.8	47	2.00	17.7
101	36.6	19.2	14.1	42	2.10	15.6
102	37.0	27.9	16.5	45	3.30	14.5
103	40.2	18.2	14.6	35	3.30	15.4
104	44.7	18.5	15.6	36	1.30	17.0
105	36.2	21.8	16.1	45	2.80	13.8
106	35.5	19.4	14.6	37	2.00	13.7
107	35.8	20.2	15.2	36	2.40	14.8
108	37.5	13.9	14.6	35	1.80	12.6
109	34.4	29.7	18.0	47	6.00	13.8
110	38.8	10.0	8.7	24	1.80	13.0
111	34.6	21.3	13.8	47	3.60	13.8
112	34.3	23.2	15.1	47	2.80	11.9
113	43.6	15.2	12.4	37	1.40	15.8
114	53.4	11.0	8.5	30	1.30	18.2
115	45.2	16.0	9.9	34	2.00	15.3
116	58.8	7.9	6.5	40	1.00	18.4
117	59.8	12.8	9.1	38	1.60	20.3
118	40.5	18.8	14.9	40	1.60	16.0
119	42.1	19.7	15.2	39	2.20	14.7
120	38.1	14.3	14.9	40	1.30	15.7
121	37.4	24.8	13.7	35	3.70	13.9
122	43.6	26.2	13.2	40	3.00	14.9
123	50.0	9.4	7.1	29	1.90	17.0
124	38.1	17.2	11.7	30	2.40	16.9
125	42.1	17.8	10.5	29	2.20	16.3
126	40.0	9.6	7.6	22	1.70	13.5
127	55.3	12.8	7.3	35	1.50	18.8
128	47.2	10.3	7.9	26	1.20	17.0
129	47.6	8.0	7.1	26	1.20	15.6
130	55.4	15.4	8.6	23	1.50	18.6
131	46.4	9.9	6.7	25	2.20	13.6
132	46.1	15.4	9.9	36	1.90	16.9

Quadro I (conclusão)

Número do registo	Casca (%) X <sub>1</sub>	D. A. P. (cm) X <sub>2</sub>	Altura (m) X <sub>3</sub>	Idade (anos) X <sub>4</sub>	Largura de copa (m) X <sub>5</sub>	Espessura de casca (%) X <sub>6</sub>
133	44.1	9.6	8.6	25	1.20	14.6
134	45.7	10.6	7.5	25	1.40	15.1
135	40.2	15.2	8.1	24	2.00	13.2
136	40.0	7.4	8.0	22	1.00	13.5
137	43.3	11.2	0.8	24	2.00	13.5
138	40.0	15.9	9.9	25	2.80	12.3
139	45.6	16.6	9.6	26	2.20	15.7
140	38.2	8.2	8.8	25	1.30	14.6
141	33.7	12.8	11.3	28	1.80	13.0
142	28.1	18.4	12.8	25	2.10	12.5
143	30.4	11.6	11.7	19	2.00	10.3
144	26.3	9.7	11.7	17	1.70	11.3
145	34.9	10.4	10.0	21	1.80	11.5
146	50.5	25.4	13.0	40	3.10	17.5
147	41.8	16.6	12.3	36	1.20	16.0
148	55.4	13.0	10.4	27	1.30	20.8
149	41.6	23.0	14.5	21	2.90	15.0
150	44.6	12.5	11.3	36	1.00	21.2
151	35.7	11.4	12.4	34	1.10	14.9
152	36.1	22.2	16.8	40	3.00	13.1
153	47.1	15.2	11.7	33	2.50	18.4
154	29.2	34.8	22.4	98	4.00	11.5
155	28.5	36.0	19.4	70	5.60	12.3
156	53.7	15.6	11.0	47	1.70	20.8
157	43.2	19.0	12.7	60	2.20	16.8
158	32.4	22.4	18.2	64	2.10	13.8
159	52.4	8.3	6.8	35	1.30	15.7
160	27.4	24.2	16.4	78	4.00	11.4
161	36.7	29.8	17.6	86	4.40	13.6
162	48.0	15.6	11.3	49	2.50	17.3
163	42.2	17.8	11.8	48	2.30	14.6
164	39.1	29.4	17.4	86	3.50	15.6
165	23.5	22.9	17.2	85	3.70	10.9
166	48.3	15.4	10.4	42	1.80	18.8
167	36.4	20.5	13.9	41	2.10	13.9
168	32.1	28.6	16.8	38	4.40	12.6
169	31.0	25.6	15.0	25	4.50	10.9
170	40.0	15.0	14.9	37	1.40	14.7
171	33.3	23.5	17.6	40	2.90	12.5
172	38.5	21.2	15.4	40	2.10	14.8
173	36.5	30.5	18.6	43	3.60	15.2
174	34.7	26.6	17.3	42	3.60	12.8
175	38.4	19.5	15.0	34	2.10	14.3
176	35.7	22.0	15.0	40	3.70	13.2
177	42.1	20.5	13.0	42	3.00	18.3
178	53.1	13.5	8.5	29	2.30	18.2
179	36.9	14.0	10.7	29	1.50	14.3
180	30.0	8.6	10.8	20	1.00	9.8
181	36.7	10.6	11.0	26	1.50	13.2
182	29.5	25.2	17.4	27	3.50	11.5
183	23.1	32.3	19.9	40	5.60	8.5
184	38.9	5.8	7.3	21	1.00	10.3
185	43.5	10.4	7.9	21	1.80	13.5
186	36.2	13.4	9.5	22	2.40	13.4
187	46.9	8.2	7.1	21	1.10	14.6
188	50.0	5.6	5.2	18	1.10	13.4
189	35.7	17.4	11.0	35	3.00	12.1
190	37.9	24.4	14.7	34	3.20	14.3
191	44.3	13.8	10.3	32	1.40	17.4
192	42.1	15.1	9.4	49	2.20	14.6
193	34.3	28.0	11.2	43	5.00	10.0
194	42.0	20.2	10.4	45	3.10	14.9
195	38.2	14.7	10.0	33	2.30	12.9
196	34.5	42.5	14.9	57	6.80	12.0
197	29.2	35.4	17.0	44	5.70	11.1
198	36.0	25.1	15.7	40	3.60	14.5

## Contributions to the Study of the Percentage of Bark in the «Pinus pinaster», Sol.

### Summary

The investigation of the agents determining the variation of bark thickness in Cluster Pine (*Pinus pinaster*, Sol.) is of the greatest importance, this species being, within the genus *Pinus*, the one, that presents the highest percentage of bark in the volume, the values ranging from 15 to 60 %.

The causes of so great a variation are due to many factors, some of which inherent to the tree, the others of ecological and cultural nature.

The first mentioned factors only have been treated, but the interference of the latter in the variation of the percent of the bark was indirectly recognized, as the morphological characters of the tree are closely associated with the ecological and cultural features of the stand.

The object of the present study has been to provide a table for estimating the percent of bark of the Cluster Pine and for that purpose, the analysis of the existing relations between different factors inherent to the tree, and the thickness of the bark, has been carried out.

Statistical methods of linear correlation were applied and the coefficients of single correlation between the different variables were calculated and all subsequent research was based upon these data.

The net regressions lines and the linear partial regressions were determined and their graphical interpretation was figured out so as to trace the nature of the variations in the percentage of bark in comparison with the factors considered.

The coefficients of partial correlation were calculated for the appreciation of the relative value of each factor tested, the reciprocity existing between them having been duly considered. The coefficients of multiple correlation were also calculated so as to determine the percentage of variation due to the independent factors considered together.

Notwithstanding the conclusions arrived at in this first part of the study, it was thought convenient to have recourse to the analysis of variance of the coefficients of multiple correlation, in order to obtain more precise indication of the factors to be made use of for the organization of the table.

The rectilinear or curvilinear nature of the regression lines was investigated by means of the analysis of variance (1).

A regression equation was determined and the table of the percentage of bark was constructed, based on the total height of tree, the age, and the percent of bark thickness at d. b. h.

The data for the present study were obtained from 198 trees, aged 17 to 109 years from several Cluster Pine stands at the State Forest of Fôja.

The characters examined were the following:

Percentage of bark.....	(X <sub>1</sub> )
D. b. h.....	(X <sub>2</sub> )
Total height of tree.....	(X <sub>3</sub> )
Age.....	(X <sub>4</sub> )
Crown size.....	(X <sub>5</sub> )
Percentage of bark thickness at d. b. h. o/bark.....	(X <sub>6</sub> )

(1) Fisher: *Test for Non-Linearity of Regression*.

## Referências bibliográficas

- BELYEA, H. C. — *Forest Measurement* (1931).  
EZEKIEL, M. — *Methods of Correlation Analysis* (1930).  
FISHER, R. A. — *Statistical Methods of Research Workers* (1930).  
GOULDEN, C. H. — *Methods of Statistical Analysis* (1936).  
NILSEN, J. — *Beregning av furens barkmasse* (Meddelelse N. R. 17 fra Vastlandets Forstlige, 1934).  
LOVE, H. H. — *Application of Statistical Methods to Agricultural Research* (1936).  
PIMENTEL, C. A. S. — *Os Nossos Pinheiros* (1910).  
REI, A. A. C. — *Coeficientes de Forma* (Relatório Final do Curso de Engenheiro Silvicultor, 1928).  
SNEDCOR, G. W. — *Statistical Methods* (1938).  
TEMUDO, J. C. F. — *Percentagem de casca no Pinheiro Bravo* (Relatório Final do Curso de Engenheiro Silvicultor, 1938).  
TIPPET, L. H. C. — *Methods of Statistics* (1937).